Published in final edited form as:

Int J Pediatr Otorhinolaryngol. 2019 January ; 116: 65–69. doi:10.1016/j.ijporl.2018.10.004.

Clinical Validation and Reproducibility of Endoscopic Airway Measurement in Pediatric Aerodigestive Evaluation

Christian R. Francom^{1,2}, Cameron A. Best³, Ryan G. Eaton⁴, Victoria Pepper⁵, Amanda J. Onwuka⁶, Christopher K. Breuer^{3,7}, Meredith N. Merz Lind, MD^{1,2}, Jonathan M. Grischkan^{1,2,6}, and Tendy Chiang^{1,2,6}

¹Department of Otolaryngology-Head and Neck Surgery, The Ohio State University, Columbus, OH, USA

²Department of Pediatric Otolaryngology-Head and Neck Surgery, Nationwide Children's Hospital, Columbus, OH, USA

³Center for Tissue Engineering, The Research Institute at Nationwide Children's Hospital, Columbus, OH, USA

⁴The Ohio State University College of Medicine, Columbus, OH, USA

⁵Department of Pediatric Surgery, Loma Linda University School of Medicine, Loma Linda, CA, USA

⁶Center for Surgical Outcomes Research, The Research Institute at Nationwide Children's Hospital, Columbus, OH, USA

⁷Department of Pediatric Surgery, Nationwide Children's Hospital, Columbus, OH, USA

Abstract

Background: Endoscopic airway measurement (EAM) combines optical endoscopic instruments with open source image processing to accurately obtain airway dimensions. Preclinical models have demonstrated EAM as an accurate technique of airway measurement with the added advantage of characterizing multilevel stenosis, non-circular lesions, and distal obstruction. The aim of this prospective clinical study was to compare EAM to airway measurements obtained from endotracheal tube approximation (ETTA) during pediatric aerodigestive evaluation and to evaluate reproducibility of EAM across practitioners.

Send correspondence to: Tendy Chiang, MD, Department of Pediatric Otolaryngology Head and Neck Surgery, Nationwide Children's Hospital, 700 Children's Drive, Columbus, OH, 43205, USA, tendy.chiang@nationwidechildrens.org.

Conflict of Interest Statement: The authors have no additional funding, financial relationships, or conflicts of interests to disclose.

Portions of this research were presented at the Annual Meeting for the American Broncho-Esophagological Association in San Diego, California (USA) from April 26–28, 2017 and the Annual Meeting for the American Society of Pediatric Otolaryngology in Austin, Texas (USA) from May 18–21, 2017.

Level of Evidence: Level II: Level 2 evidence is prospective observational research with an experimental design (i.e., "cohort study")

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Methods: Thirty-seven pediatric patients undergoing routine microlaryngoscopy and bronchoscopy at a single tertiary care children's hospital were prospectively recruited. Patients undergoing emergent procedures were excluded. Two blinded reviewers performed airway measurements using ImageJ (NIH) as previously described and average values were compared to ETTA measurements. Additional EAMs were obtained from an *ex vivo* airway model by 28 separate clinicians and were analyzed by the same reviewers to evaluate reproducibility.

Results: EAM and ETTA measurements were themselves significantly different ($p = 0.0003$); however, the average absolute difference between the two methods was small (Mean: 0.5 mm, 95% CI: -2.6 to 1.6 mm). There were notable differences between raters such that estimates of raters with more experience were more similar to ETTA. Despite observed differences between EAM and ETTA, endoscopic airway measurement was highly correlated with ETTA ($p = 0.0002$, Spearman $r = 0.4185$), and strong agreement was observed (Bias: -0.4974 ± 1.083 mm, 95% LOA: -2.62 to 1.625 mm).

Conclusion: Clinical use of EAM is a valid and precise approach for quantification of airway luminal dimensions. This method may provide advantages over traditional ETTAs for evaluation of asymmetric airway morphology in the pediatric population.

Keywords

Quantitative Endoscopy; Airway Stenosis; Image Processing; Bronchoscopy; Endoscopic Airway Measurement; ImageJ

1. Introduction

Important in the evaluation and management of laryngotracheal pathology is the ability to accurately measure airway dimensions, as this information can drive interventions. Subjective sizing of the airway is known to be inaccurate and unreliable, even when performed by experienced practitioners [1]. In 1994, Cotton and Myer published a landmark study describing an objective method of airway grading using endotracheal tube approximation (ETTA) to size the airway [2]. This technique has been the gold standard for more than 20 years. There are situations, however, where ETTA may be inaccurate or difficult to perform. Non-circular stenosis, multilevel stenoses or distal tracheal pathology in the airway often prove challenging to measure with ETTA.

To address the limitations of ETTA, a number of methods have been described in the literature that seek to improve the accuracy of airway measurement. In some cases, non-standard bronchoscopic equipment is used to measure airway dimensions [3,4,5]. These techniques have not been widely adopted, presumably as a result of the need for new surgical instruments or other equipment. Some described methods utilize the video imaging data obtained through bronchoscopy to reconstruct computer-generated models of the airway, which then allows for endoluminal dimensional analysis [6,7,8]. While these techniques have each shown promising results, they have not become widely adopted given the requirement of novel computer software, specialized equipment, and expert training to execute data analysis.

Alternatively, radiographic measurements can be acquired with computed tomography (CT), magnetic resonance imaging (MRI), or ultrasound (US), but require additional expense to the patient and may be inaccurate due to variances in point selection, secretions and/or motion artifact. The ionizing radiation required for CT scans continues to be a concern, especially in the pediatric population. Further, there is increasing concern regarding the anesthetic administration that may be required to minimize movement and anxiety for either CT or MRI. While these tools may be used to size the airway when they are obtained for alternative indications, endoscopic evaluation with ETТА will continue to be the gold-standard in the absence of a readily available method that offers advantages in terms of accuracy, ease of use, and safety.

One method that addresses many of these limitations is endoscopic airway measurement (EAM) which utilizes standard bronchoscopic equipment and an open source image processing software platform (Image J v1.49, NIH). The technique is readily available without requiring new equipment or complicated programming. We previously demonstrated that EAM is a reliable strategy to characterize airway dimensions both in an *ex vivo* tracheal simulation and *in vivo* using an ovine model [9]. The aim of this study was to prove the equivalence of EAM to the current gold standard in noncomplex airways. The secondary aim of this study was to assess reproducibility and ease of use across practitioners. Demonstrating equivalence and ease of use in standard airway evaluation supports the feasibility of obtaining objective measurements in endoscopic imaging in a multitude of otolaryngologic applications including complex airway pathology, endoscopic sinus and skull base surgery.

2. Material and Methods

Human Research Ethics Statement

Institutional review board (IRB) approval was obtained from the appropriate governing committee at Nationwide Children's Hospital prior to initiation of the study (IRB # IRB15–00789). All patient data were securely stored and discarded when necessary per institutional policy. Pediatric patients undergoing routine microlaryngoscopy and bronchoscopy at a single tertiary care children's hospital were prospectively recruited to participate during preoperative evaluation. All participants had the right to refuse enrollment without a change in the care they were to receive. Informed consent was obtained from the legal guardians of all patients enrolled, and assent documentation was obtained from all patients between nine and eighteen years of age. Patients undergoing emergent procedures and those undergoing foreign body removal were excluded. The attending physician performing the procedure also had the right to withdraw the patient from the study if there was any intraoperative concern regarding potential risk or harm to the child for any reason. No patient was withdrawn from the study due to these concerns.

Bronchoscopy and Image Acquisition

We conducted a prospective nonrandomized single-site clinical study of 37 patients undergoing standard direct laryngoscopy and full bronchoscopic evaluation including ETТА for airway measurement as previously described by Cotton and Myer [2]. Patients were

placed in a supine position with gentle neck extension. The head and body axle and head neck yield angle was maintained in a clinically consistent position for each procedure. Briefly, an endotracheal tube is passed through the larynx and a Valsalva maneuver is performed to gradually increase the airway pressures up to 25 cm H₂O as the surgeon assesses for an air leak. If there is an air leak at pressures between 10 and 25 cm H₂O, then it is determined that the outer diameter of that given endotracheal tube approximates the inner diameter of the airway. The amount of time it took to perform ETTA was recorded. Each patient then underwent airway measurement using EAM as previously described [9]. The steps of EAM included the following: an optical forceps was attached to a Hopkin's rod telescope at the standard manufactured distance from the end of the lens. An image was captured in the airway at the level of stenosis or at the level of the subglottis if no stenosis were present. A second image was captured with a ruler at the level of the forceps to provide a calibration scale. Images were then de-identified and securely saved for later processing. Patient information and demographics including age, gender, and indications for bronchoscopy were recorded. The remainder of the patient's preoperative and postoperative care was standard for the procedures being performed. The total operative time was recorded.

Image Analysis

Two blinded raters then performed airway measurements from the intraoperative images using Image J (v1.49, NIH). One rater is a general surgeon with limited experience with image processing software (VP), and the second rater is a doctoral student with a decade of basic science expertise with image processing software (CAB). Each rater first established a scale of distance-to-pixel-ratio from the calibration image. This distance-to-pixel-ratio was then applied to the associated bronchoscopic image. The freeform tool in Image J (v1.49, NIH) was used to calculate intraluminal airway dimensions by tracing the internal circumference of the airway and correlating with the distance-to-pixel-ratio.

Statistical Analysis

ETTA is a useful clinical instrument for the quantitative assessment of the airway, and EAM can also be used to evaluate the airway quantitatively [10,11]. Luminal area measurements from each rater were transformed to reflect luminal diameter, allowing for direct comparison to ETTA via the following formula: $\left(\text{Diameter} = 2\sqrt{\frac{\text{area}}{\pi}} \right)$. EAM and ETTA measurements were considered paired if they were derived from the same level of the airway in the same patient.

To determine if there were any statistically significant differences between EAM and ETTA measurements, we employed two strategies. The first involved the averaging of EAM scores across raters. ETTA and EAM averages were compared with a Wilcoxon matched-pairs signed rank test because measurements failed the D'Agostino & Pearson normality test. The second strategy was to estimate a multilevel regression model. This strategy did not require averaging airway measurements across raters, which were occasionally quite different. Rather, this model assessed the degree to which measurement technique and rater caused variation in the airway measurement. We employed the Kenward-Roger degrees of freedom

method to adjust for small sample bias and assessed Likelihood Ratio tests to identify the best fitting model. Our final model was a three-level model which accounted for multiple observations nested within each person nested within each Rater. This model included fixed effects for evaluation method, random intercepts at the patient level and rater level, and a random coefficient for evaluation method.

Finally, a nonparametric Spearman correlation was computed between EAM and ETTA, and prompted a Bland Altman agreement analysis to determine the interchangeability of the two techniques. Type I error was controlled at $\alpha = 0.05$ and two-sided p values < 0.05 were considered statistically significant. Statistical analysis was performed with GraphPad Prism 7.0 and SAS Enterprise Guide.

Reproducibility Study

The secondary objective of this study was to evaluate reproducibility of EAMs. Twenty-eight clinicians enrolled in an airway training program at our institution were recruited and enrolled in this study. Participants ranged in experience from medical students to attending physicians, and the specialties represented amongst the participants included Otolaryngology, Emergency Medicine, Anesthesia, General Surgery, and Intensive Care. A brief tutorial was given to the participants to instruct them on bronchoscopy and EAM image acquisition. The subjects were able to practice bronchoscopy on an advanced infant patient simulator (SimBaby® Laerdal Medical) with the model of subglottic stenosis⁹. The SimBaby is a simulator, and thus an artificial design, however it normalizes the conditions in order to evaluate ease of use of EAM. Once the participants felt comfortable with the equipment, each participant individually performed a direct laryngoscopy and bronchoscopy on the Sim Baby® and captured an image for EAM without assistance. Two blinded reviewers, the same mentioned above, subsequently calculated the airway dimensions from the test images using Image J (v1.49, NIH). The area of subglottic stenosis in the Sim Baby® was measured with both an endoscopically placed ruler and with endotracheal tube approximation which both provided the same result of an inner diameter measuring 8.5 mm. This value was used to evaluate the accuracy of measurements for each participant group using a one-sample t-test. Subsequently, we assessed differences in mean measurements by provider experience using one-way ANOVA and Tukey's HSD for groups with unequal observations.

3. Results

Study Participants

A total of 37 patients were enrolled in the EAM validation study (Table 1). The average age was 3.9 years (SD 4.23) and 67% were male. Indications for bronchoscopy included evaluation of an existing tracheostomy tube (22%), recurrent croup (14%), recurrent aspiration (11%), laryngomalacia (11%), evaluation of known subglottic stenosis (8%), stridor (5%), tracheomalacia (5%), vocal cord paralysis (5%), and other (19%). The average operative time was 13.84 minutes (SD 8.52), with image acquisition taking on average 47.54 seconds (SD 6.72) and ETTA taking 124 seconds (SD 81). There were no adverse outcomes or other complications to any of the patients enrolled in this study.

Comparison of EAM and ETTA Measurements

In the first analysis strategy, mean EAMs (across raters) were compared to traditional ETTA (Fig. 1a). Pairwise analysis revealed a significant difference between the two modalities ($p = 0.0003$). The multilevel analysis demonstrated similar results. Fixed effects for the evaluation method suggested that there was a significant difference in average airway measurements between EAM and ETTA ($p=0.003$). EAM assessments were, on average, smaller than that of ETTA. Random intercepts for patient and rater were significant ($p=0.003$ and $p=0.005$, respectively), as well as the random coefficient for evaluation method ($p=0.003$).

Despite the difference between the measurements, EAM and ETTA yielded estimates that were quite similar. The absolute difference between the two methods was 0.5 mm with a 95% Limit of Agreement (LOA) between -2.6 to 1.6 mm. Further, EAM and ETTA were significantly correlated (Fig. 1b, $p = 0.0002$, Spearman $r = 0.4185$), and a Bland-Altman agreement analysis (difference vs. mean) indicated a strong agreement between the two modalities (Fig. 1c, Bias: -0.4971 ± 1.083 mm, 95% Limits of Agreement: -2.62 to 1.625 mm).

Reproducibility

Twenty-eight clinicians captured an image for EAM without assistance and two raters calculated the airway dimensions from the test images using Image J. Images from 3 clinicians were excluded given poor image quality. The known inner diameter of the SimBaby® trachea was 8.5 mm, and the average measurement provided by study participants was 8.9 (Range: 7.8, 9.6). The average difference between raters was 0.19, however this difference was not significantly different ($p=0.22$). Clinicians with any bronchoscopy experience had airway measurements similar to the known dimensions of the simulation trachea; however, clinicians with no previous bronchoscopy experience tended to have overestimated airway measurements (Fig. 2, $p=0.0018$). Similarly, clinicians with 5 or more post-graduate years had airway measurements similar to the known dimensions of the simulation trachea; however, individuals with less experience tended to measure larger ($p\text{-value}_{PGY<3} = 0.0003$ and $p\text{-value}_{PGY<5} = 0.02$). Finally, measurements of clinicians with >30 direct laryngoscopies were the same as known dimensions of the simulation trachea; however, this differed for individuals with less laryngoscopy experience ($p\text{-value}_{<16} = 0.003$ and $p\text{-value}_{<31} = 0.001$).

4. Discussion

Endoscopic airway measurement has been described as a viable strategy for characterizing airway dimensions when compared to direct and radiographic measurements in a large animal (ovine) model of tissue engineered tracheal graft (TETG) implantation [12]. The current study pursued the clinical translation of this approach by comparing EAM to the gold standard of objective airway measurement, ETTA, in the pediatric population and assessing the reproducibility of measurement.

Our results demonstrate that there is a significant correlation between the ETTA and EAM methods in a pediatric population. The absolute difference between the two methods was 0.5 mm (5.9%), suggesting that EAM can be used interchangeably with ETTA sizing because the mean difference is clinically insignificant. Further, this study demonstrates that the technique is not only safe, but that it adds minimal time in the operating room. There was no intraoperative or postoperative morbidity within the patient cohort. As the mean operative time to perform this technique was less than one minute, it was faster on average than the current standard.

EAM is a sensitive metric for which lack of experience can impact the precision of image capture. The quality of the image is dependent on position of instrument in an area only a few mm in diameter, and as such, slight variations in image capture can have drastic differences in airway assessment. Our data demonstrates that EAM is reproducible between clinicians of various training and experience levels, although some experience may be beneficial in order to accurately perform and reproduce EAM. This finding was illustrated in both the EAM validation study and the reproducibility study where the raters and clinicians with more expertise yielded airway measurements that were more similar to ETTA. The ETTA has maintained its popularity because it is a straightforward and easy to perform method; however, this study provides preliminary evidence that EAM also has these advantages. Interestingly, the degree of over-estimation appears to correlate with provider experience.

This method is not intended to replace endotracheal tube approximation as described by Myer et al. [2], but was borne out of necessity. Although the goal this study was to demonstrate equivalence in noncomplex airways, EAM also has the potential to measure irregular, multilevel and distal segments of stenosis that were not measurable with standard techniques and EAM may also be a method used to validate airway measurement derived by dynamic volume CT angiography. The rationale for including only non-complicated airways in this study was that the EAM for less complicated airways should be more similar to the ETT, and that in more complicated airways (non-circular for example), we would expect more of a difference between ETT and EAM with EAM measurements potentially being more accurate. Further, because EAM has the potential to increase procedure time, we wanted to minimize harm to patients that have higher degrees of stenosis or more complex indications for the bronchoscopy. Indeed, additional research is necessary to better understand EAM performance in more complex airways and also, for which patients is EAM most similar to ETTA.

Another limitation of this study was that additional variability was introduced by having different individuals perform the image capture from those measuring the airway on the image. In clinical application, the provider who performs the bronchoscopy would also take the image and assess the airway on the image in real time. Having designated image readers was an attempt to introduce rigor to the study and also meet limitations in study resources that prohibited clinicians from also receiving training to read, calibrate and assess images. However, this decision likely introduced additional error into the study as some degree of subjectivity is involved in scaling the distance-to-pixel-ratio from the calibration image. As such, we hypothesize that in a truly applied clinical scenario, there would be less variability

in the EAM and ETTA measurements than demonstrated in this study, and more time required on the part of the clinician to download and read images.

Clinical and translational adoption of EAM is governed in part by its validity and the reproducibility of measurements, which were each evaluated in this current study. This technique is readily accessible, as it requires only standard bronchoscopy equipment and a free software package (Image J). Additional research will further elucidate the advantages and limitations of this approach, however, our findings support that endoscopic airway measurement provides the ability to reproducibly quantify airway dimensions in the management of airway stenosis and this method should be

5. Conclusion

EAM provides a reliable technique for airway characterization, and the methods themselves are reproducible across clinicians, though more experience is associated with improved performance. Our findings suggest that this method of airway luminal measurement is at the very least interchangeable with traditional ETTA approximation and should be considered for use in pediatric aerodigestive evaluation.

References:

1. Murgu S, Colt H. Subjective assessment using still bronchoscopy images misclassifies airway narrowing in laryngotracheal stenosis. *Interactive CardioVascular and Thoracic Surgery*. 2013;16(5): 655–660. [PubMed: 23407694]
2. Myer CM, O'Connor DM, Cotton RT. Proposed grading system for subglottic stenosis based on endotracheal tube sizes. *Ann Otol Rhinol Laryngol*. 1994;103(4):319–323. [PubMed: 8154776]
3. Nouraei SAR, McPartlin DW, Nouraei SM et al. Objective Sizing of Upper Airway Stenosis. *Laryngoscope*. 2006;116(1):12–17. [PubMed: 16481801]
4. Hayashi A, Takanashi S, Tsushima T, Denpoya J, Okumura K, Hirota K. New method for quantitative assessment of airway calibre using a stereovision fiberoptic bronchoscopy. *British Journal of Anaesthesia*. 2012;108(3):512–516. [PubMed: 22201182]
5. Sharma GK, Foulad A, & Verma SP A Novel Device for Measurement of Subglottic Stenosis in 3 Dimensions During Suspension Laryngoscopy. *JAMA Otolaryngol Head Neck Surg*. 2015;14(4): 377–381.
6. Bromwich M, Murugappan M, Wilging JP. Adding the third dimension – A new tool for constructing 3D models of the airway from 2D bronchoscopic video. *International Journal of Pediatric Otorhinolaryngology*. 2009;73(9):1202–1207.
7. Calloway HE, Kimbell JS, Davis SD et al. Comparison of Endoscopic versus 3D CT Derived Airway Measurements. *Laryngoscope*. 2013;123(9):2136–2141.
8. Meisner EM, Hager GD, Ishman SL, Brown D, Tunkel DE, Ishii M. Anatomical Reconstruction of Pediatric Airways from Endoscopic Images: A Pilot Study of the Accuracy of Quantitative Endoscopy. *Laryngoscope*. 2013;123(11):2800–2887.
9. Pepper VK, Francom C, Best CA et al. Objective characterization of airway dimensions using image processing. *International Journal of Pediatric Otorhinolaryngology*. 2016;91:108–114. [PubMed: 27863622]
10. Pandian V, Hutchinson CT, Schiavi AJ, Feller-Kopman DJ, Haut ER, Parsons NA, Lin JS, Gorbakkin C, Angamuthu PG, Miller CR, Mirski MA, Bhatti NI, Yarmus LB. Predicting the need for nonstandard tracheostomy tubes in critically ill patients. *Journal of Critical Care*. 2017; 37:173–178. [PubMed: 27756050]

11. Hodnett BL, Simons JP, Riera KM, Mehta DK, & Maguire RC. Objective endoscopic findings in patients with recurrent croup: 10-year retrospective analysis. *International journal of pediatric otorhinolaryngology*. 2015;79(12):2343–2347. [PubMed: 26574171]
12. Pepper VK, Onwuka EA, Best CA et al. Endoscopic management of tissue-engineered tracheal graft stenosis in an ovine model. *Laryngoscope*. 2017; 127(10): 2219–2224. [PubMed: 28349659]

Figure 1:

(a) comparison, (b) correlation, and (c) agreement of EAM and ETTA. A significant difference between methods was found by Wilcoxon matched-pairs signed rank test ($p = 0.0003$), and the measurements were significantly correlated ($p = 0.0002$, Spearman $r = 0.4185$, 95% CI: 0.2051 to 0.5939, $R^2 = 0.2683$). Bland-Altman analysis indicated strong agreement between measurements derived from both methods of airway sizing (Bias: -0.4974 ± 1.083 mm, 95% LOA: -2.62 to 1.625 mm). Dashed lines indicate 95% LOA.

Figure 2:

EAM measurement by clinician years of training. One sample t-test comparing EAM measurements derived from images based on years of training to the known airway stenosis diameter of 8.5 mm (dashed line). There significant difference between the known diameter and measurements from the PGY 1–2 (** $p = 0.0003$) and PGY 3–4 (* $p = 0.0189$) groups. There was no difference between the known diameter and measurements from the groups with most experience (PGY 5–6, 7–9, fellow / attending groups).

Table 1:

Patient Demographics and Indications for Bronchoscopy

	n (%)
Gender	
Male	25 (67%)
Female	12 (33%)
Age (SD)	3.9 (+/- 4.23)
Indication for procedure	
Existing tracheostomy	8 (22%)
Recurrent croup	5 (14%)
Aspiration	4 (11%)
Laryngomalacia	4 (11%)
Subglottic stenosis	3 (8%)
Stridor	2 (5%)
Tracheomalacia	2 (5%)
Vocal cord paralysis	2 (5%)
Other	7 (19%)